Appendix C: Digital Data Capture

This appendix brings together material from various sections of the *Guide*, in expanded form, to provide a detailed description of how analog information is converted into digital data in various media types. While for many purposes this level of technical detail may be more than is needed, a basic understanding of the principles involved can be useful in evaluating the appropriateness of certain types of equipment, determining when digitization is likely to yield good results (or not), and understanding why certain kinds of conversion can result in data loss or degradation. Specific recommendations for formats, settings, and how to get the best results from different kinds of materials are addressed in the main sections of the *Guide*; the goal here is to provide a more detailed explanation of the basic principles involved.

General Principles

Analog and digital data are fundamentally different: where analog information is generally smooth and continuous, digital information consists of discrete chunks, and where analog information bears a direct and non-arbitrary relationship to what it represents, digital information is captured using formal codes that have only an arbitrary and indirect relationship to the source. Thus while an analog image, for instance, consists of continuously varying colors and shading, a digital image consists of a set of individual dots or pixels, each recording the color intensity and other information at a given point. Although the specific kinds of information vary from medium to medium—sound waves, light intensities, colors—this basic difference remains a constant.

Conversion from analog to digital thus requires that the continuous analog information be sampled and measured, and then recorded in digital format. There are several basic factors which govern this process and which determine the quality of the digital result.

The first of these is the density of data being captured from the analog original: in effect, how often the original is sampled per unit of time (in the case of video and audio) or area (in the case of images and video). For digital audio, the higher the sampling rate, the smoother the transitions between the individual packets of sound, to the point where, with modern digital audio, they cannot be detected by the human ear. A low sampling rate results in clipping, the audio equivalent of jerky animation. For digital images, the higher the sampling rate (i.e. resolution), the smoother and less pixellated the image appears, and the more it can be magnified before its granularity becomes visible.

The second factor at work is the amount of information that is recorded in each sample. Individual pixels in an image may contain very little information—at the most minimal, they may take only one binary digit to express on versus off, black and white—or they may take 32 bits to express millions of possible colors. Large sample size may be used, as in digital images, to capture nuance, finer shadings of difference between values. They may also be used to express a wider total range, as in the case of digital audio, where a higher frequency response means that the recording can capture a greater range of frequencies, with higher highs and lower lows.

Both sampling frequency (or resolution) and sample size (frequency response, bit-depth) involve a trade-off of data quality and file size. It is clear that the more frequently you sample, and the more information you capture in each sample, the larger your file size will be, and the more costly to create, transmit, store, and preserve. Decisions about digital data capture are thus not simply a matter of achieving the highest possible quality, but rather of determining the quality level that will represent the original adequately, given your needs. Various sections of

the *Guide* explore these considerations in more depth.

The remainder of this appendix describes how these principles apply in detail in particular digital media.

Digital Audio and Video Capture

In analog audio recording, a plucked string (for example) vibrates the air around it. These airwaves in turn vibrate a small membrane in a microphone and the membrane translates those vibrations into fluctuating electronic voltages. During recording to tape, these voltages charge magnetic particles on the tape, which when played back will duplicate the original voltages, and hence the original sound. Recording moving images works similarly, except that instead of air vibrating a membrane, fluctuating light strikes an electronic receptor that changes those fluctuations into voltages.

Sound pressure waveforms and other analog signals vary continuously; they change from instant to instant, and as they change between two values, they go through all the values in between. Analog recordings represent real world sounds and images that have been translated into continually changing electronic voltages. Digital recording converts the analog wave into a stream of numbers and records the numbers instead of the wave. The conversion to digital is achieved using a device called an analog-to-digital converter (ADC). To play back the music, the stream of numbers is converted back to an analog wave by a digital-to-analog converter (DAC). The result is a recording with very high fidelity (very high similarity between the original signal and the reproduced signal) and perfect reproduction (the recording sounds the same every single time you play it no matter how many times you play it).

When a sound wave is sampled using an analog-to-digital converter, two variables must be controlled. The first is the sampling rate, which controls how many samples of sound are taken per second. The second is the sampling precision, which controls how many different gradations (quantization levels) are possible when taking the sample. The fidelity of the reproduced wave can never be as accurate as the analog original; the difference between the analog signal and the closet sample value is known as quantization error. This error is reduced by increasing both the sampling rate and the sampling precision. As the sampling rate and quantization levels increase, so does perceived sound quality.

In digital representation, the same varying voltages are sampled or measured at a specific rate, (e.g. 48,000 times a second or 48 kHz). The sample value is a number equal to the signal amplitude at the sampling instant. The frequency response of the digital audio file is slightly less than half the sampling rate (Nyquist Theorem). Because of sampling, a digital signal is segmented into steps that define the overall frequency response of the signal. A signal sampled at 48 kHz has a wider frequency response than one sampled at 44.1 kHz. These samples are represented by bits (0's and 1's) that can be processed and recorded. The more bits a sample contains, the better the picture or sound quality (e.g., 10-bit is better than 8-bit). A good digital signal will have a high number of samples (e.g., a high sampling rate) and a high number of bits (quantizing). Digital to digital processing is lossless and produces perfect copies or clones, because the digital information can be copied with complete exactness, unlike analog voltages. High bit-depth is also result in much-increased dynamic range and lower quantization noise.

Ideally, each sampled amplitude value must exactly equal the true signal amplitude at the sampling instant. ADCs do not achieve this level of perfection. Normally, a fixed number of bits (binary digits) is used to represent a sample value. Therefore, the infinite set of values possible in the analog signal is not available for the samples. In fact, if there are R bits in each sample, exactly 2R sample values are possible. For high-fidelity applications, such as archival copies of analog recordings, 24 bits per sample, or a so-called 24 bit resolution, should be used. The difference between the analog signal and the closest sample value is known as quantization error. Since it can be regarded as noise added to an otherwise perfect sample value, it is also often called quantization noise. 24-bit digital audio has negligible amounts of quantization noise.

Digital Image Capture

Digital image capture divides the image into a grid of tiny regions, each of which is represented by a digital value which records color information. The resolution of the image indicates how densely packed these regions are and is the most familiar measure of image quality. However, in addition to resolution you need to consider the bit-depth, the amount of information recorded for each region and hence the possible range of tonal values. Scanners record tonal values in digital images in one of three general ways: black and white, grayscale, and color. In black and white image capture, each pixel in the digital image is represented as either black or white (on or off). In 8-bit grayscale capture, where each sample is expressed using 8 bits of information (for 256 possible values) the tonal values in the original are recorded with a much larger palette that includes not only black and white, but also 254 intermediate shades of gray. In 24-bit color scanning, the tonal values in the original are reproduced from combinations of red, green, and blue (RGB) with palettes representing up to 16.7 million colors.

Digital Text Capture

Although it may seem odd to discuss digital text in this context, there are some important, if indirect parallels between the principles described above and those that govern digital text capture. Clearly in capturing digital text one does not sample the original in the same way that one samples audio or images. However, the process of text capture does involve choices about the level of granularity at which the digital representation will operate. In capturing a 20th-century printed text, for instance, a range of different "data densities" is possible: a simple transcription of the actual letters and spaces printed on the page; a higher-order transcription which also represents the nature of textual units such as paragraphs and headings; an even more dense transcription which also adds inferential information such as keywords or metrical data. Other possibilities arise in texts that have different kinds of internal granularity. In the case of a medieval manuscript, one might create a transcription that captures the graphemes—the individual characters—of the text but does not distinguish between different forms of the same letter (for instance, short and long s). Or one might capture these different letter forms, or even distinguish between swashed and unswashed characters. One might also choose to capture variations in spacing between letters, lines of text, and text components, or variations in letter size, or changes in handwriting, or any one of a number of possibly meaningful distinctions.

These distinctions, and the choice of whether or not to capture them, are the equivalent of sampling rates and bit-depth: they govern the amount of information which the digital file records about the analog source, and the resulting amount of nuance that is possible in reusing and processing the digital file.

---

valid xhtml 1.1
abp~03/03